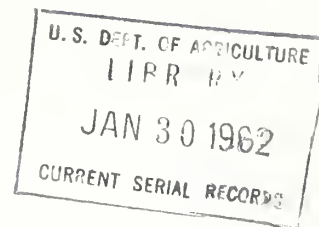# Historic, archived document

Do not assume content reflects current
scientific knowledge, policies, or practices.

UNITED STATES DEPARTMENT OF AGRICULTURE
Agricultural Marketing Service
Cotton Division

PROGRAM FOR
MULTIPLE CORRELATION AND REGRESSION EQUATIONS
One Dependent and Up to Five Independent Variables
on the IBM 650 Digital Computer

By Elgin G. Fry, Cotton Marketing Specialist

ACKNOWLEDGMENTS

PROGRAM FOR
MULTIPLE CORRELATION AND REGRESSION EQUATIONS
One Dependent and Up to Five Independent Variables
on the IBM 650 Digital Computer

This program uses simple correlations, standard deviations, and means to calculate the following statistical values: multiple correlation coefficient, variance, regression equation coefficients, beta values, and partial correlation coefficients. Standard errors for the estimating equation, regression coefficients, and beta values are also calculated. Any method desired may be used to obtain the simple correlations, means, and standard deviations. While the program is designed for use by the Cotton Division, Agricultural Marketing Service, United States Department of Agriculture in statistical analysis of cotton properties, it can be readily adapted to any type multiple correlation problem when the prescribed input format is followed.

The program uses a control column to determine how many different sets of multiple correlations are required in each problem. All of the above statistical values are calculated and punched for a minimum of one dependent and two independent variables. If more than two independent variables are included in the problem, a complete set of values for one dependent and two independent variables will be calculated and punched before calculating the values for one dependent and three independent variables. This process will continue up through one dependent and five independent variables; therefore, a total of four complete problems can be made from one set of cards by use of the control column (one dependent with two independents, one dependent with three independents, one dependent with four independents, and one dependent with five independents). The computer calculates all of the above values through one dependent and five independent variables in approximately 20 seconds.

The program is designed to give the same sign to the partial correlation coefficient as its respective b value.

A standard 10 digit-8 word control panel is used with the 533 Card Read Punch machine.

Input

The means and standard deviations necessary in the calculations must be placed in the low-order position of their respective 10 digit words with the decimal position between the fifth and sixth columns in each word (xxxxx.xxxxx). Decimals are never punched. Each mean and standard deviation is limited to five significant figures to the right of the decimal position. All of the simple correlation coefficients must be placed in the high-order position of their respective words with the decimal position at the beginning of each word (.xxxxxxxxxx). Each simple correlation coefficient can be carried out to ten places if so desired, but should contain at least four or five significant figures. Any remaining portion of the 10 digit word which is not used must be filled with zeros.

Minus signs must be punched in the units position for all negative correlation coefficients. Plus signs are not normally punched since each word will be read as positive through the proper wiring of the 533 control panel if no sign is indicated in the card.

Any identification code desired may be placed in columns 3 through 10 (word 1) of each input card. Columns 1 and 2 of each input card is the card number.

The attached format (Exhibit A) shows the layout of data used for special test cards. The number of independent variables must be shown in column 11 of card 3. Three cards are used for two independent variables;

and six cards for five independent variables. If all six cards in a set are not used, the unused cards must not be included. If all words are not used in the cards containing the means and standard deviations of the independent variables, then the remaining words should be filled with zeros.
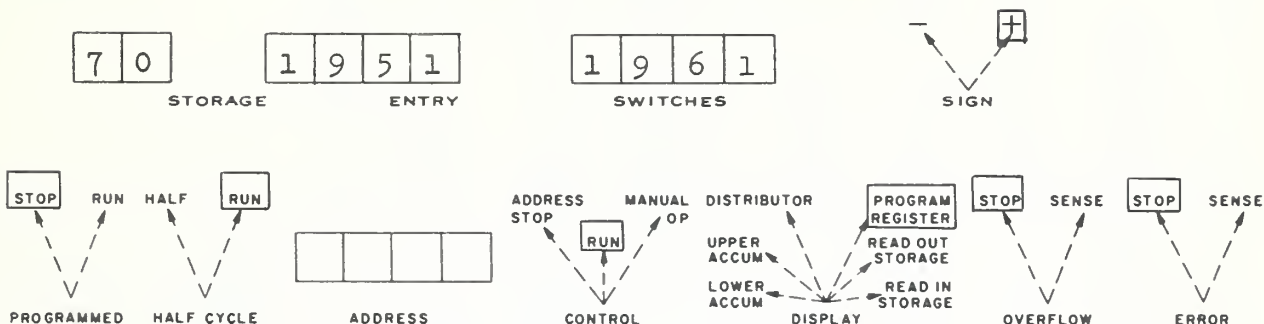
Output

A simplified method of recording the output data is indicated by the attached output sheet (Exhibit B). The answers were derived from the data shown in Exhibit A. A 407 control panel has been specially wired to give the exact spacings for each value; however, a standard 80 column-8 word panel may be used. Each answer in the output cards has five significant figures to the right of the decimal point. In every case, the decimal point is located between the fifth and sixth columns in each word. The identification code given in card 1 (columns 3-10) of the input data is placed in columns 3-10 in all output cards. Columns 1 and 2 give the card number for the output data.

Copies of the program and test cards may be obtained by requesting Library Number 1.0.026 from the Statistical Reporting Service, United States Department of Agriculture, Washington 25, D. C.

OPERATOR'S CHECK LIST

IBM 650 PROGRAM OPERATING INSTRUCTIONS

I. Initial Console Setting as shown above.
   A. Normal Starting Procedure: Computer Reset; Program Start.
   B. Special Instructions: _____

II. Card Input - Output (533 or 537)

READ FEED

NO. OF CARDS	FILE DESCRIPTION
	PROGRAM DECK
	TEST CARDS
	DATA CARDS

PUNCH FEED

CARD FORM
8-10 DIGIT WORD CARDS
ANY COLOR

CONTROL PANELS

READ 8-10 DIGIT WORDS
PUNCH 8-10 DIGIT WORDS
RSU AND PSU ONLY

III. Program Stops and Required Action:

STOP ADDRESS	MESSAGE - EXPLANATION - ACTION
	STOP 1: MISPUNCHED CARD OR BLANK COLUMN.
	(a) CORRECT CARD AND INSERT IN PROPER PLACE.
	(b) START ALL CARDS AGAIN FOR SET CONTAINING ERROR CARD
	(01, 02, 03, 04, 05, 06, - IF ALL SIX CARDS ARE USED)
	(c) COMPUTER RESET AND PROGRAM START.
	STOP 2: OVERFLOW
	(a) CHECK COLUMN 11 OF CARD 3 FOR PROPER NUMBER.
	(b) IF COLUMN 11 IS CORRECT, REMOVE SET FROM DATA DECK.
	(c) COMPUTER RESET AND PROGRAM START.
	(d) BEGIN WITH FIRST CARD OF NEXT SET.

EXHIBIT A

INPUT DATA FOR MULTIPLE CORRELATIONS

R₁ _____ thru _____ Year _____ Study _____ No. Obs. _____

Dep. Var. (1)	Var. (2)	Var. (3)	Var. (4)	Var. (5)	Var. (6)
_____	_____	_____	_____	_____	_____

Instructions to Keypunch Operators:
Punch only lines completely filled out. Ignore decimals.
Minus signs punched in units position when indicated.

SAMPLE OF
INPUT DATA FOR TEST CARDS

WORD 1	WORD 2	WORD 3	WORD 4	WORD 5	WORD 6	WORD 7	WORD 8
Ident.	S.D. Var. 2	S.D. Var. 3	S.D. Var. 4	S.D. Var. 5	S.D. Var. 6		
0103210260	00004.83450	00000.59050	00000.91210	00000.04380	00000.34080	0000000000	0000000000
	Mean Var. 2	Mean Var. 3	Mean Var. 4	Mean Var. 5	Mean Var. 6		
0203210260	00073.31440	00008.46380	00003.08260	00001.02280	00004.28820	0000000000	0000000000
	S.D. Dep. Var.					Mean Dep. Var.	OBS
0303210260	50007.79300	.1963400000	.0269600000	.0995800000	0000000000	00116.61250	2710000000
	r14	r12	r13	r23			
0403210260	.2465300000̅	.3288800000̅	.3546000000̅	0000000000	0000000000	0000000000	0000000000
	r15	r24	r34	r45	r56		
0503210260	.5599600000	.0420800000	.2454300000̅	.0147400000̅	0000000000	0000000000	0000000000
	r16	r25	r35	r46			
0603210260	.0966800000	.0543400000	.0920800000̅	.2196500000̅	.2051500000	0000000000	0000000000
		r26	r36	r46	r56		

INSTRUCTIONS:

1. The number of Independent Variables will be placed in control column of Card 3 (column 11).
2. If all fields for S.D. and Means are not used, these fields must be filled with ten zeros.
3. The minus (-) sign will be shown above the units position in each word if the correlation coefficient is minus.
4. The number of observations in original study will be placed in columns 71-73 of Card 3 (OBS).

MULTIPLE CORRELATIONS AND REGRESSION EQUATIONS

R _____ thru _____ Year _____ Study _____ No. Items _____

Card No.	Study No.	Org.	Color	Code	Dep. Var.	Var. 2	Var. 3	Var. 4	Var. 5	Var. 6		
1	3	2	1	260		$R_{1.23}$ .19647+	$R^2_{1.23}$ .03860+	$S_{1.23}$ 7.64111+	° 92.66358+	b_2 .31528+	b_3 .09858+	$se\ b_2$ .09691+
2	3	2	1	260			$se\ b_3$ .79303+	B_2 .19559+	B_3 .00747+	$se\ B_2$ .06012+	$se\ B_3$ .06009+	
3	3	2	1	260			$r_{13.2}$ .00707+					$r_{13.3}$ .19468+
4	3	2	1	260		$R_{1.234}$ .28205+	$R^2_{1.234}$ .07955+	$S_{1.234}$ 7.47660+	° 114.94292+	b_2 .20627+	b_3 .87973−	b_4 1.94872−
5	3	2	1	260		$se\ b_2$ .10018+	$se\ b_3$ .82813+	$se\ b_4$ .56510+	B_2 .12796+	B_3 .06666+	B_4 .22808−	
6	3	2	1	260		$se\ B_2$ .06215+	$se\ B_3$ .06275+	$se\ B_4$ .06614+	$r_{12.34}$ .12494+	$r_{13.24}$ .06481−	$r_{14.23}$ .20637−	
7	3	2	1	260		$R_{1.2345}$ .62307+	$R^2_{1.2345}$ .38822+	$S_{1.2345}$ 6.09537+	° 7.62210−	b_2 .17251+	b_3 1.29822+	b_4 1.43471−
8	3	2	1	260		b_5 102.68078+		$se\ b_2$ .08173+	$se\ b_3$ .70064+	$se\ b_4$ .46274+	$se\ b_5$ 8.84096+	B_2 .10702+
9	3	2	1	260		B_3 .09837+	$se\ B_3$ .04969+	B_4 .16792−	B_5 .57711+	$se\ B_2$ .05309+	$se\ B_4$ .05416+	
10	3	2	1	260				$r_{12.345}$ .12802+	$r_{13.245}$ .11261+	$r_{14.235}$ .18628−	$r_{15.234}$ .57909+	
11	3	2	1	260		$R_{1.23456}$ .62583+	$R^2_{1.23456}$ .39166+	$S_{1.23456}$ 6.07823+	° 1.72052−	b_2 .16953+	b_3 1.17773+	b_4 1.58269−
12	3	2	1	260		b_5 104.51338+	b_6 1.41820−	$se\ b_2$ .08169+	$se\ b_3$ .70698+	$se\ b_4$ .47795+	$se\ b_5$ 8.96017+	
13	3	2	1	260		$se\ b_6$ 1.15683+	B_2 .10517+	B_3 .08924+	B_4 .18524−	B_5 .58741+	B_6 .06202−	
14	3	2	1	260		$se\ B_2$ .05068+	$se\ B_3$ .05357+	$se\ B_4$ .05594+	$se\ B_5$ .05036+	$se\ B_6$ .05059+		
15	3	2	1	260		$r_{13.2456}$ .10154+	$r_{14.2356}$ .19890−	$r_{15.2346}$ .58155+	$r_{16.2345}$ .07490−			$r_{12.3456}$ .12613+

Posted by _____   Date _____